Early Journal Content on JSTOR, Free to Anyone in the World

This article is one of nearly 500,000 scholarly works digitized and made freely available to everyone in the world by JSTOR.

Known as the Early Journal Content, this set of works include research articles, news, letters, and other writings published in more than 200 of the oldest leading academic journals. The works date from the mid-seventeenth to the early twentieth centuries.

We encourage people to read and share the Early Journal Content openly and to tell others that this resource exists. People may post this content online or redistribute in any way for non-commercial purposes.

Read more about Early Journal Content at http://about.jstor.org/participate-jstor/individuals/early-journal-content.

JSTOR is a digital library of academic journals, books, and primary source objects. JSTOR helps people discover, use, and build upon a wide range of content through a powerful research and teaching platform, and preserves this content for future generations. JSTOR is part of ITHAKA, a not-for-profit organization that also includes Ithaka S+R and Portico. For more information about JSTOR, please contact support@jstor.org.

all. They are not the units as they have often been made. It is from this fact that the end of the sentence may control the beginning as well as the beginning control the end. No part can be isolated from the whole, and in the whole must be considered the antecedent thought, large masses of earlier experiences, the present environment, particularly the social environment, the emotional as well as the intellectual condition of the individual at the time of speaking.

W. B. PILLSBURY.

UNIVERSITY OF MICHIGAN.

INDIVIDUAL DIFFERENCES IN BELIEF, MEASURED AND EXPRESSED BY DEGREES OF CONFIDENCE

THE present study only reports in part results obtained from a series of experiments with common beliefs. Among the classes of material used in related series of experiments are horizontal lines, handwriting, and tones. Complete results of all of these experiments will soon be published. Our diverse kinds of material were employed in an experimental endeavor to give quantitative meaning to the degrees of confidence a person experiences while making various kinds of judgments.

The present investigation throws considerable light on individual differences, especially when a scale for measuring "degrees of confidence" is used. The experiments herein reported afford answers to three problems:

Problems.—1. We may assume that objects, with as great differences between them as one wishes, will be discriminated by every one with absolute certainty. It will also be granted that objects, with sufficiently less differences between them, will be discriminated by every one with something less than absolute surety. The problem then arises: What is the correlation between an individual's "degrees of confidence," and the "differences discriminated" by him? As these differences become larger, will his increasing confidence-strength in their discrimination be nicely graduated to the increasing objective differences? What differences are here found among individuals? What are the probable causes of such individual differences as appear?

2. Does the correlation, in the case of an individual, between "degrees of confidence" and "differences discriminated," bear a significant relation to a correlation (in the case of the same individual) of his judgments of discrimination and those of a large group of persons? That is, does an individual's high or low approximation to a social standard bear any definable relation to his own correlation of *confidence* and *discriminated difference?*

3. Can subjective confidence be defined and quantitatively measured in terms of objective differences discriminated? This problem has a special significance with reference to the highly subjective material (beliefs) of our experiment.

Material.—Twenty-five beliefs, or statements (as we prefer to call them), were selected in the following manner. Dr. E. K. Strong had recast and revised sets of beliefs, used in previous investigations: he had a usable set of about fifty. From this set the author drew twenty-five by chance. Three of those drawn, which Strong had found to be uncertain in their behavior, were exchanged for three others. Our twenty-five statements, with their identification numbers, follow.

2. ''There exists an all-wise Creator of the world.''
3. '' Easy come, easy go.''
4. '' Traffic in alcoholic beverages, except for medicinal purposes, should be nationally prohibited.''
5. '' Only the good die young.''
9. '' For the benefit of the individual an eight-hour working day should be strictly enforced.''
10. '' Breaking a mirror brings on seven years of misfortune.''
11. '' See a pin and pick it up, and all the day you'll have good luck.''
12. '' The happiest life for woman is the married life.''
13. '' The maternal instinct is strong in most women.''
15. '' Geo. Washington was a real person.''
17. '' Virgil wrote the ' Æneid.' ''
21. '' Our Republic will endure another hundred years.''
22. '' $2 + 2 = 4$.''
23. '' $2 + 4 = 7$.''
24. '' The moon is larger than Jupiter.''
26. '' The earth is practically round.''
33. '' Christianity is losing its influence in this country.''
34. '' The people of the world are becoming better.''
35. '' It would benefit the majority of young men to attend college.''
37. '' Women should have the vote.''
39. '' Music more nearly approaches the divine than does poetry.''
46. '' Death ends personal existence.''
47. '' The whale swallowed Jonah.''
48. '' Man is free to work out his own salvation.''
50. ''Blessed are the meek, for they shall inherit the earth.''

Conduct of the Experiments.—The subjects for these experiments were thirty-five students of Barnard College, members of Professor Hollingworth's classes in experimental psychology; all had had considerable experience in similar kinds of work.

In the first experiments, the girls arranged the twenty-five statements in an order of belief-strength. The Order of Merit Method was employed. Most of the subjects had two trials, separated by about a week. In no case were they aware that a second trial was to follow the first. Each student was given a set of the statements, with the following directions:

BELIEF

> Arrange these in an order; first select the one you *believe the most.* Second, select the one you *believe next strongly,* and so on, until the one you *believe the least* is selected last.

The second experiment was conducted by the elaborate Method of Paired Comparisons; that is, every statement was compared with every other, by each of the thirty-five subjects. Thus, with the twenty-five statements there were *three hundred* comparisons or "tests." Each subject was provided with blanks, on which were printed the identification numbers of the tests in order, from 1 to 300. Duplicates of the statements were provided, so that each subject had ready at hand the "three-hundred-comparison-test" material,—three hundred cards, to each of which was fastened two statements to be compared. To avoid a possible space error, the positions of identical statements on successive cards were changed. For example, the card for Test No. 1 had statement 2 on the left side and statement 3 on the right; but the card for Test No. 2 had statement 4 on the left and statement 2 on the right side. The usual way of conducting an experiment by the Method of Paired Comparisons was followed.[1] The written instructions were as follows:

> These statements are to be compared according to the *strength* of your *belief* in them.
>
> 1. In each of the 300 comparison-tests, to the right of each test-number (on blank provided), record the number of the statement in which you *believe* the *more.*
>
> 2. Then,—using the scale for Degree of Confidence,—also record your *degree of confidence* in your decision.

The scale of confidence employed was derived from similar scales of other workers. The form of our present scale, while embodying all of the merits of earlier scales, has this advantage: a definition of each degree of confidence in three different ways, together with at least *a priori* equal differences between each of the four degrees. Even before beginning the experiments, the author did not think that the different degrees of confidence would mean the same to any two subjects, or that any one person would be able, in the actual experience of an experiment, to hold (subjectively) to the *a priori* set equal degree-differences of the scale. Those who took part in the experiments were instructed to hold to the defined conditions of the scale as closely as possible. Any person, after a very little experience with the scale, came to attach his own subjective and precise meanings to each degree of confidence, and consistently to maintain his own set differences between each step (unequal as these differences often were).

---

[1] Titchener, "Experimental Psychology, Student's Manual, Qualitative," pages 92–95.

The scale follows:

DEGREE OF CONFIDENCE

Use a scale of *four degrees,* as follows:

0 = A mere guess. Lack of confidence.—Absence of any surety.
1/3 = 33 1/3 per cent. surety.—Somewhat certain.—Some confidence of being right, but faintly so.
2/3 = 66 2/3 per cent. surety.—Very certain.—Reasonably sure. Considerable confidence of being right.
3/3 = 100 per cent. surety.—Absolutely certain.—As strong a confidence as one would have under these particular conditions.

Calculations of Results.—The results of the second experiment, that by the Method of Paired Comparisons, were calculated in the usual way. In the course of the 300 tests, each statement appeared equally often, and thus had a chance of being chosen in preference to any of the others. The statement that was chosen or "believed the more" the largest number of times by any one subject was ranked first in her *scale of belief-strength* (such we may call it, in so far as the 25 statements used are concerned); that statement which was preferred the next to the greatest number of times by any subject was ranked second in her *scale of belief-strength;* and so on, with the times preferred of the 25 statements, the one preferred the least number of times (or not at all) being ranked last or twenty-fifth in the individual's *scale of belief-strength.* Thus, by an elaborate series of tests, a scale of *belief-strength* was obtained for each of the thirty-five subjects. We then turned back to the original records, on the filled-out blanks of 300 tests each. In this case we were not concerned with the particular statement chosen in each case, but with the *degree of confidence* attached to that choice or decision (see "written instructions," above). Now, what is meant by saying that a certain degree of confidence is "attached to a decision"? There is a discrimination of difference, with reference to any two statements involved in any one of the 300 tests. The confidence is not attached to the belief chosen, for had that same chosen belief been compared with some other one (more like or more different), the confidence would have been different in many cases. Thus, "$2+2=4$" being compared with "$2+4=7$," the confidence will be "3/3," or "3," in most cases; but if "$2+2=4$" is compared with "The earth is practically round," may not the confidence be often less than absolute? As a matter of fact, it generally was. Therefore, if the confidence is not attached to the statement chosen, we must say it is attached to both of any two compared, that is, to the difference that a person discriminates between them. To ascertain what any such difference discriminated (or "paired comparison rank difference," as we shall hereafter call it) is, we have only to refer to our subject's scale of belief-strength (described above). Consequently, the procedure was as follows. A tally sheet was prepared for each girl, with P. C. (paired comparison, rank differences ranging from 0 to 24, the greatest possible difference in any individual's scale of belief-strength. The beginning of such a sheet follows:

SUBJECT 1. TALLY SHEET

P. C. Rank Differences	Degrees of Confidence	Average Confidence
0	0000 111 22	0.78
0.5		
1.0	000 11111	0.63
1.5	0000 11 22	0.75
2.0	0000 11111111111111 22	0.90
2.5	000 1 2 3	1.00
3.0	00 111112222 3	1.33

The Degree of Confidence (in one and all of the 300 tests) was first noted (it being 0, 1, 2, or 3). This Confidence was put on the tally sheet opposite that P. C. Rank Difference, which was the difference between the two statements compared, as ascertained by subtracting the ranks of the two statements on the subject's scale of belief-strength. We thus treated the confidence in each of the 300 tests, for any one individual. (The results of each person were computed without reference to any other.) Since several pairs of statements would have the same Rank Difference, in any one's scale of belief-strength, there are generally several entries (on the tally sheet) under Degrees of Confidence for each P. C. Rank Difference. Thus (above) for Difference 0 there are nine entries, for Difference 1.0, eight, etc.

Thus any one of our tally sheets has two significant columns, the first and the last (see above): (1)a column of *rank differences,* and (2) a column of *average confidences,* one for each rank difference. We next correlated these average confidences with the rank differences, or "differences discriminated" (see Problems, 1, above). This was done for each subject separately, without reference to any one else's records. In making the correlations, the following formula was used:

$$r = 1 - \frac{6(\Sigma d^2)}{n(n^2 - 1)}.$$

The results are given in Table I.

In addition to the above correlations of each subject, without reference to any of the others, the *median* confidence, for all of the subjects, for each rank difference, was calculated. These *median confidences* and *rank differences* were correlated, with a resulting coefficient of $+ .956$.

TABLE I

Correlations of Degrees of Confidence and Differences Discriminated
These Coefficients of Correlation are All Positive

S	*r*	*S*	*r*	*S*	*r*	*S*	*r*	*S*	*r*
7	.705	31	.833	5	.863	34	.896	29	.920
3	.750	27	.846	14	.864	28	.900	32	.921
21	.791	24	.851	13	.872	30	.903	22	.927
10	.809	11	.853	25	.879	18	.905	2	.928
9	.816	15	.854	16	.881	1	.906	17	.933
35	.817	20	.854	12	.887	4	.911	33	.948
19	.828	6	.854	23	.891	26	.911	8	.961

"S" refers to the subjects, by identification numbers.

With reference to the second of our problems (see above), a "social standard" or position for each of the twenty-five statements was obtained in the following manner. Seventy-four different persons judged the statements, or ranked them in order of belief-strength. Both the Order of Merit and the Paired Comparisons methods were used. There were 146 judgments concerning (or rankings of) *each* of the 25 statements,—3,650 judgments in all. Our result was a group or "social" position for each statement. This position, in the case of each statement, was the *median* of the 146 rankings of that statement. Thus we secured 25 "group medians," our "social standard."

For each of the 35 persons the rankings in the one or two trials by the Order of Merit Method were averaged with the rankings of the set of statements secured as a result of the Paired Comparisons experiment, to secure, for each subject, an average, individual position for each belief. For each observer, then, her average order of ranking of the 25 beliefs was correlated with the "group medians" (described in the above paragraph). We thus found each individual's high or low approximation to a fairly reliable social standard. For the purposes of the present report, these coefficients need not be given.

We have now secured two sets (of 35 each) of coefficients of correlation, the one last mentioned, and that of Table I. (above). Each individual has two coefficients, one measuring her "tendency towards concomitant variation" of confidence and discriminated difference, the other measuring the same relation with respect to her approximation to a social standard. To secure a quantitative answer to the second problem,—the two sets of coefficients (with measurements of the two traits in question for each of the 35 observers) were correlated. The resulting coefficient was —.035, or practically zero. That is, under the conditions of this experiment, the two traits are absolutely devoid of correlation. We are warranted in saying that an individual's nicely graduated increase of confidence to correspond to a graduated increase in her differences discriminated, will *not* enable us to make any inference whatsoever regarding the fact of her high or low approximation to the social standard of her fellows.

There yet remains the third problem: with the highly subjective material used, can subjective confidence be quantitatively measured in terms of objective differences discriminated? I believe I am able to show that it can. The reader should refer back to the part of a sample tally sheet. The procedure was to ascertain the *differences discriminated*, or *"rank differences,"* corresponding to the *median* of the *average confidence* (for each subject). In the tally sheet, these average confidences are found in the last column. Zero confidence was defined as any degree of confidence from 0 to 0.99 (thus including such figures as, "0.78," "0.63," etc.); confidence 1 was defined as any degree from 1 to 1.99 (thus including such values as, "1.00," "1.33," etc.); confidence 2 was defined as any degree from 2 to 2.99 (inclusive); and confidence 3 was the highest possible degree. Now, average confidence zero has a range of from 0.00 rank difference to difference 3.5 (not all to be seen in the part of the sample tally sheet given). In the case of the same subject, average confidence 1 has a range of from rank difference 2.5 to 6.00 (with a median at 5.00). Average confidence 2 has a range of from difference 4.00 to 15.00. Average confidence 3 has a range of from 10.00 to 24.00 (the highest difference discriminated in the case of this subject). In her case, the rank differences corresponding to the medians of each degree of average confidence were as follows: confidence 0, 1.5; confidence 1, 5.00; confidence 2, 9.00; and confidence 3, 18.25. In the same way, median rank differences for each of the four degrees of confidence, for each subject, were obtained. Thus, for each degree of confidence, we have a series of 35 (the number of subjects) median rank differences. For each series, the median and the mode were found, which values may be regarded as *objective* measures of the subjective degrees of confidence. Table II gives the facts.

TABLE II

Subjective Confidence Defined in Terms of Objective Differences Discriminated

Degrees of Confidence	Medians	Modes	Q.
0	1.50	1.50	0.50
1	4.25	3.50 and 4.00	0.88
2	10.50	11.50	1.50
3	18.00	17.75	1.13
Differences Between Confidences			
0 and 1	2.75	2.25	
1 and 2	6.25	7.75	
2 and 3	7.50	6.25	

Thus, with the persons tested, we see that, quantitatively and objectively, the steps taken in passing from any degree of confidence to the degree next higher are very unequal. A problem arises at this point, which we will not attempt to solve. Assuming the logical soundness of the statement: "That we pass by degrees from one thing to another, is no sign that the things are of the same nature,"—we may doubt whether the different degrees of confidence are of the same nature. But especially does this become a significant problem, in the light of the great inequality of differences between the degrees.

I believe that it has now been shown that confidence can be *objectively* defined. An objection may arise as to the use of the word "objectively," to refer to an individual's scale of discriminated differences (in terms of which are the medians and modes of Table II.) in connection with such subjective material as beliefs. While, it may be urged, a certain belief is chosen the greatest number of times (in the Paired Comparisons experiment) and ranked first, another is chosen a next fewer number of times and ranked second, and so on, until the one chosen least often is ranked last,—the one ranking first *could* be preferred say 24 times, the second in rank could be preferred say 23 times, but the one ranking third *could* be chosen only 17 times (provided none were preferred 18, 19, 20, 21 or 22 times), while the fourth in rank could be chosen *any* number of times—less than 17, not to mention the possibilities of several statements each being chosen the same number of times. Clearly, in such cases, an individual's scale would be anything else than "objective" (in the sense that a foot-rule, for example, is an objective scale). To show that this very legitimate objection is groundless in the present study, we tabulated the number of times that each belief (of course regardless of what particular one it might be) was chosen by each subject—in order, from largest to smallest number of times preferred. For Subject No. 1, for example, the order of times preferred is as follows: 24, 23, 22, 21, 19, 18, 18, 17, 17, 15, 13, 12, 12, 12, 11, 10, 8, 5, 5, 5, 4, 3, 3, 2, 1. It is seen that the steps of this scale are about equal, though this particular individual had more identical numbers of times preferred than most of the thirty-five. The order of times preferred of all the subjects was strikingly similar. Average times preferred for each step (of the 25 in all of the subjects' scales of discriminated differences) were computed. These averages, with their average deviations are found in Table III.

Thus our *average* scale of discriminated differences in terms of times preferred clearly shows that the material employed afforded a scale of steps very nearly equal. While this measuring-rod is not as absolutely "objective" as a foot-rule or a pair of balances, I think we are justified in using the word "objective" and in concluding that, even with such material as beliefs, subjective confidence can be measured and defined in terms of objective differences discriminated.

TABLE III

Average Times Preferred	A. D.	Average Times Preferred	A. D.
23.95	0.10	11.20	0.67
22.83	0.28	10.03	0.33
21.77	0.40	9.11	0.40
20.74	0.50	8.09	0.42
19.37	0.57	6.89	0.51
18.46	0.65	5.91	0.47
17.37	0.58	4.91	0.47
16.74	0.53	4.00	0.34
15.94	0.54	3.31	0.47
15.17	0.47	2.57	0.55
14.23	0.40	1.40	0.53
13.29	0.65	0.34	0.46
12.37	0.75		

We are now in a position to consider more carefully certain aspects of the first problem. The conclusions reached are summarized in Table I. (above). In this table we have positive proof of a very high correlation between *degree of confidence* and *difference discriminated.* As such differences become larger, increasing confidence-strength in their discrimination is nicely graduated to the increasing objective differences. Such is the accepted meaning of a coefficient of correlation. In no individual do we find a low degree of confidence for a small difference, remaining low, until the objective difference has grown to considerable proportions, and then "jumping" to a higher degree; on the contrary, the mutual relation is one that is "nicely graduated"; the tendency is towards concomitant variation.

So much for any and all of the subjects. In the statement of our problems (above), the question was also asked, "what differences are here found among individuals?" Table I. (above) shows these individual differences at a glance. The coefficients range from +.705 to +.961, with a median at +.879. While the individuals arrange themselves in a fairly continuous series, the difference between the extremes is large enough to warrant an attempt at answering the further question (set in the statement of problems): "What are the probable causes of such individual differences as appear?" At the present stage of our investigation, and with the very inadequate knowledge of individual differences in general, only a probable and tentative answer can be now given.

As a possible cause of these differences among individuals, we may consider individual scales of belief-strength (or rank differences), with which the average confidences were correlated. Perhaps different people had such scales of very unequal steps (as determined by the number of times the different beliefs were chosen in the paired comparisons experiment), that our resulting series of coefficients were thereby caused to differ (Table I.). Table III. (above), and its explanation, proves that such was by no means the case. On the contrary, the "A.D.'s" of Table III. clearly prove that the thirty-five observers had scales of belief-strength almost identical at each step.

As another cause of the individual differences of Table I., it might be thought that the subjects gave the different degrees of confidence rather different values (as objectively determined in Table II.); it will be remembered that we there measured confidence in terms of objective differences discriminated, or "rank differences." The reader, at this point, should refer back to the discussion immediately preceding Table II.

For each subject, the A.D. of the average difference between the objective rank differences for each of the four degrees of confidence (of that subject), gives a measure of *how unequal* were the steps between the four degrees—of *how different* were the objective values assigned to each degree. The larger the A.D., the greater is the inequality of the degrees with reference to each other. The smaller the A.D., the more equal are the steps from one degree to each succeeding degree. Will persons differing widely in this respect reveal a corresponding individual difference, as shown in Table I.? To obtain a quantitative answer, the thirty-five A.D.'s (one for each subject) secured were correlated with the coefficients of Table I. The resulting coefficient was —.02. Again we have failed to find a cause for our individual differences.

The fact that, by nature, some individuals are more conservative and others are more radical and confident in most of their decisions, may possibly explain the individual differences in question. From the definition and construction of our scale of confidence we may assume that a *conservative* person would, when making his judgments, set Confidence 3 at a much greater distance from Confidence 0, than an individual of more *radical* turn of mind—one inclined to be absolutely sure of most of his opinions and beliefs. Therefore (the reader should again refer back to the discussion immediately preceding Table II.), the conservative person would show the *greater* differences between the objective rank differences of confidences 0 and 3 (the extremes of the scale); and the radical individual would show the *less*. Perhaps the individual differences of Table I. may be partly due to some of the observers being more radical or more conservative (in the sense defined) than the others. To determine this —for each subject—the objective values secured for confidences 0 and 3 were subtracted. The 35 results (forming a series, measuring the trait in question for all of the subjects) were correlated with the coefficients of Table I. The resulting coefficient was + .435. This figure is not high enough to afford positive proof that we have discovered a cause of the individual differences; it does show a tendency for the more conservative persons to have the higher coefficients (Table I.). Further investigations are needed, with this and other material, that we may follow up this very possible cause of individual differences.

Summary of Results.—First problem: 1. A high degree of correlation was found between an individual's *degrees of confidence and differences discriminated:* increasing confidence-strength is nicely graduated to increasing objective differences. The coefficients of correlation range from + .705 to + .961, with a median at + .879; when the median confidences (for all of the subjects) are correlated with the differences discriminated, the resulting coefficient is + .956.

2. The range of values for the coefficients, just mentioned, show marked individual differences.

3. As to the probable causes of these individual differences: it is found,—first, that inequality of the steps in the individual scales of belief-strength can not be regarded as a cause,—second, it is seen that persons differing widely as far as giving the different degrees of confidence different objective values, do not reveal a corresponding individual difference,—third, a coefficient of + .435 shows a tendency for the more conservative persons (in the sense defined) to have the higher coefficients (Table I.), but further investigations must follow up this possible cause of individual differences.

Second problem: 4. Under the conditions of this experiment, an individual's nicely graduated increase of confidence to correspond to a graduated increase in differences discriminated, will *not* enable us to make any kind of inference regarding the fact of her high or low approximation to a social standard.

Third problem: 5. We find that subjective confidence can be quantitatively measured in terms of objective differences discriminated, even with such material as common beliefs. The amounts of difference between the four degrees of confidence are as follows: between 0 and 1, 2.75; between 1 and 2, 6.25; and between 2 and 3, 7.50 (Table II.).

George F. Williamson.

Columbia University.